DOCUMENT RESUME

ED 355 235 TM 018 486

TITLE	Gauging High Performance: How To Use NAEP To Check Progress on the National Education Goals. Report of the NAEP Technical Subgroup for the National Education Goals Panel.
INSTITUTION	National Education Goals Panel, Washington, DC.
REPORT NO	NEGP-92-01
PUB DATE	24 Jan 92
NOTE	18p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Standards; *Achievement Tests; Educational Assessment; Educational Improvement; Educational Objectives; Elementary Secondary Education; *Evaluation Methods; Guidelines; High Achievement; *Longitudinal Studies; *Performance; Student Evaluation; Test Construction; Test Items; *Test Use
IDENTIFIERS	Curriculum Alignment; *National Assessment of Educational Progress; *National Education Goals 1990; National Education Goals Panel; Performance Based Evaluation; Progress Reporting

ABSTRACT

The National Assessment of Educational Progress (NAEP) is one part of a comprehensive assessment system that is evolving in the United States. In the view of the NAEP Technical Subgroup for the National Education Goals Panel, the NAEP should be an independent, statistically rigorous longitudinal measure of educational attainment in the essential subject matters. To make the NAEP a real part of educational future improvement, the following directions are recommended: (1) alignment of NAEP frameworks and test objectives to national content standards; (2) reporting NAEP results at state and national levels; (3) quality and breadth of items and tasks in each subject matter area; (4) subject matter and grade levels roughly as in the current NAEP; (5) unambiguous and carefully defined achievement levels; (6) development of performance assessment tasks; (7) assessment in a 3-year cycle; and (8) solid research to follow up interesting information coming out of the NAEP. The NAEP must be reauthorized periodically. The state-by-state NAEP, in particular, is a temporary enhancement that will disappear without congressional action. The National Education Goals Panel can be a champion for this reauthorization. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED355238

NATIONAL EDUCATION GOALS PANEL

GAUGING HIGH PERFORMANCE:

How to Use NAEP to Check Progress on the National Education Goals

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
CHARLES J. WALKER

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Report of the NAEP Technical Subgroup for the National Education Goals Panel

January 24, 1992

92-01

BEST COPY AVAILABLE

NATIONAL EDUCATION GOALS PANEL
MEMBERS

GOVERNORS

Carroll A. Campbell, Jr., South Carolina, Chairman (1991–1992)
John Ashcroft, Missouri
Evan Bayh, Indiana
Terry Branstad, Iowa
Howard Dean, Vermont
Benjamin Nelson, Nebraska
Barbara Roberts, Oregon
Roy Romer, Colorado, Past Chairman (1990–1991)

MEMBERS OF THE ADMINISTRATION

Lamar Alexander, Secretary of Education
Roger B. Porter, Assistant to the President for Economic and Domestic Policy

MEMBERS OF CONGRESS

**Representative Richard Gephardt
**Representative Robert Michel
**Senator George Mitchell
**Senator Bob Dole

**Designee May Be Named

Members of the NAEP Technical Subgroup

Richard P. Mills, Chair
Vermont Commissioner of Education

R. Darrell Bock
Professor, University of Chicago

David Cohen
Professor, Michigan State University

Paul Copes
Superintendent of Schools, Bloomfield, Connecticut

Robert Glaser
Professor and Director of LRDC
University of Pittsburgh

David Haynes
Associate Teacher for Humanities
Saturn School of Tomorrow, St, Paul, Minnesota

Mark Musick
President, Southern Regional Education Board

Suzanne Triplett
Assistant Superintendent for Research and Development
North Carolina Department of Public Instruction

Leland Yee
Commissioner, San Francisco Board of Education

Panel Staff

Edward Fuentes
Martin Orland

We acknowledge the comments of the following colleagues who reviewed earlier drafts: Checker Finn, Terry Peterson, Gordon Ambach and Lauren Resnick.

GAUGING HIGH PERFORMANCE

How to use NAEP to check progress on the National Education Goals

The report of the NAEP Technical Subgroup for the National Education Goals Panel

Like a racing team working through the night to prepare their formula one car, the National Education Goals Panel is bent on high performance, not basic transportation. That is the meaning of the National Education Goals. Now the Panel must pick the right package of gauges to measure performance so that we can act when we need to. There will have to be more than one gauge: some can be adapted from what we have in hand, and others have to be invented.

Our immediate starting point was the report of the National Education Goals Panel Resource Group on Student Achievement (Goal 3) which pointed to the National Assessment of Educational Progress (NAEP) as the best available data for monitoring progress toward the national goals. We concur in this judgment. As the bibliographical notes indicate, our discussions also rest on much earlier analysis. Our charge was "to develop recommendations for monitoring progress toward the National Education goals using the National Assessment of Educational Progress. We were to "produce a report that outlines an interim and long-term data collection plan," including:

. alignment of NAEP frameworks and test objectives to national content standards

. reporting at the national and state levels

. quality and breadth of items and tasks in each subject matter area

. subject matter areas and grade levels to be assessed

. establishment of achievement levels

. use of performance assessment tasks

. periodicity of the assessment

. measures of other relevant variables such as instructional practices

This paper appears amid a complex discussion about change in education, and about measuring performance at local, state and national levels. Multiple choice testing is pervasive in the schools, and Chapter 1 is one of the factors that keeps it going, although discussions about the reauthorization of Chapter 1 are building support for a change to other forms of

assessment. Many educators would welcome an alternative because they don't believe that multiple choice tests measure what is important for students to know and be able to do.

Several states are committed to building new performance based approaches to assessment, and teachers are part of the design teams. A large consortium, the New Standards Project, proposes a national system of examinations stressing portfolios, performances and projects. The Council of Chief State School Officers is also organizing states into consortia to create assessments connected to national standards. The first state by state comparisons on mathematics are in hand from the National Assessment of Educational Progress, and a second report is on the way. The National Goals Panel has released the first national report card on the goals, and many states and local communities have their own report cards. The Panel worked with the Congress to create the National Council on Education Standards and Testing and the Council's recent report is a thoughtful guide through the complex policy issues concerning national assessment and standards. The President set a challenging timetable for a national examination system, which has induced many to think about how much can be done how fast.

This discussion will continue until the nation reaches consensus on the need for a set of high performance standards, a way to create and maintain them, a connected strategy for measuring and reporting performance in relation to those standards at local, state and national levels, and a robust strategy to use what we learn to boost performance. Our contribution to that debate concerns a piece of the emerging system, which is NAEP.

Where does NAEP fit?

NAEP is one part of a comprehensive assessment system that is evolving. Other parts will include local and state assessments, international comparisons, and explicit links between assessments, curriculum, new teaching practices, and professional development. The connecting link in all of this will be national content standards that are high and that have wide support.

Our views on NAEP's place in this can be stated briefly at the outset. NAEP should be an independent, statistically rigorous longitudinal measure of educational attainment in the essential subject matters. It should be the very best it can be in this role. It should report progress on national goals publicly at the national and state levels by assessing representative samples of students. National education goals three and four are the real focus. NAEP should also collect selected student and school variables needed to monitor changes in the population of students and schools. NAEP will probably remain the primary source of performance data on Goals three and four until the end of the decade.

NAEP is evolving from two decades of experience and we note several themes pointing to its future direction that we support strongly:

. **Closer connection to national standards.** The 1990 state by state NAEP took shape as the National Council of Teachers of Mathematics created its framework of standards, and NAEP attempted to incorporate some of the intent of those standards.

. **State by state reporting.** The 1990 state by state NAEP introduced a powerful aid to the nation's thinking about education: data to enable states to track their own progress. The power of these data makes it more important for NAEP to embrace the emerging national content standards. Were this not to happen, some members of our panel caution, the effect of NAEP would be retrograde, not progressive.

. **Proficiency levels.** Proficiency levels, which define the performance expected of students, enable the public to understand in concrete terms what the nation's students know and can do. This is vital information.

. **Faster reporting.** While it took 16 months to report the results of the 1990 state NAEP, the 1992 results will be available in ten months or less.

. **More performance questions.** Approximately two thirds of the NAEP items are multiple choice. However, NAEP is gradually moving toward more performance oriented kinds of assessment by using open ended questions. While open ended questions are not all all that is commonly meant by performance assessment today, they are a start.

. **Concern about the future and the past.** NAEP has a long series of reports to its credit, but the national goals represent a watershed event for the assessment of educational progress in this country. It is now important to check changes from our current situation without losing the ability to follow trends from the seventies. NAEP is changing and that is appropriate as long as NAEP preserves its role as a stable, accurate set of statistics.

A vision from five years out

Assessment is changing fast. What was standard practice five years ago is unacceptable today. We tried to envision what assessment might look like five years from now, in order to provide a context for our recommendations. There are many possible futures, and our recommendations do not depend on this one:

In 1996, NAEP is still the major source of national data on performance, but other elements of the system are operating. Consensus on national content standards emerged slowly in the first half of the decade. NAEP evolved to approximate the new standards as they came on line. Huge numbers of people participated in discussions of what all students should know and be able to do, first in the states and then nation-wide. Most people acknowledge that creating and communicating those content standards was the crowning achievement of the reform effort.

A consortium of states, which now represents nearly half of the nation's enrollment, has been building a national system of examinations based on portfolios and other alternative assessments. Their work is far from done, but they have pilot tested portfolios, and have proved that large scale performance assessment works.

A striking feature of the assessment landscape in 1996 is the acceptance of classroom teachers as designers of assessments. This involvement, particularly at local and state levels, produced great diversity in the assessment tasks and questions. Teacher involvement in assessment has provoked an explosive, and costly, demand for professional development.

In schools, performance assessment predominates but there were initial rough spots over how to guarantee consistency and manage time requirements. Computer image processing to manage portfolios is still a dream for most schools but many among the first generation of schools to embrace performance assessment have gone this route. Multiple choice testing has found a much smaller niche, but it appears to be secure now that even the most ardent assessment reformers have acknowledged that some achievement data can be gleaned efficiently in this way. However, the nation has outgrown its decades long dependence on standardized multiple choice tests.

Many states filled in a basic repertoire of assessments in the core subjects to meet legislative pressures to report results across the curriculum. State assessments include many NAEP items. Some states accomplished this through joint ventures by sharing design costs. There is continuous debate, both in the National Education Standards and Assessments Council and elsewhere, over what matches the national standards and what does not.

Reports on the national goals have appeared each year, and NAEP is a centerpiece. States and many communities have similar reports. The general public is more sophisticated about performance data.

Each year, NAEP reports on yet another area of the curriculum. One of the developments that encouraged greater understanding was agreement by mid decade that results from a whole set of assessments should come out in concert just before the opening of school each year. This raised the visibility of performance data and showed respect for the local policy making cycle which was needed to drive change.

The nation is making some progress, but not nearly as quickly as optimists had hoped. The annual drumbeat of reporting and discussion of those reports has convinced most people of just how difficult change will be. Concern about equity of results is a part of every debate on performance.

Now in 1996, it is clear that the national assessment again needs a strategic mid course correction. But the overall direction selected in 1992 still seems right, and most Americans are far more interested in boosting the results than in changing the testing system.

Making the NAEP part of that vision real

In reality, of course, it will work out differently. We will move both faster and slower. While we think we see the major forces, we have no doubt missed some important ones. But this vision focuses our attention on the main question: What must we do now? This is what we recommend:

1. Alignment of NAEP frameworks and national content standards

The national content standards should drive NAEP. The National Assessment Governing Board's (NAGB) apparent willingness to do this is visible in its efforts to alter NAEP to incorporate the NCTM standards as they have emerged since 1988. Everyone will have to manage through a period of uncertainty in the next few years because national standards development will impose on an already defined test development schedule. Today, the NAEP mathematics test does not fully match the NCTM standards. However, about half of the items of the eighth grade mathematics tests were judged by mathematics education experts to be related to cross-cutting themes of the NCTM curriculum and evaluation standards.*

The science assessment is another case in point. The National Academy of Science is developing content and assessment standards for 1994, and NAEP is creating a science assessment to be administered and the results reported in 1994. These two efforts are not on the same cycle. The NAEP science assessment would have to be postponed to 1996 to fully capture the Academy's work. Nevertheless, collaboration over the next 18 months could cause these two efforts to influence each other, to the benefit of both.

Because NAEP has historically operated in the absence of national standards reflecting what we expect students to know and be able to do, it has tended to measure what is taught and learned rather than a national consensus on what should be. However as we envision a time when national standards become more prominent, we would expect this to change. And in fact, NAEP has attempted to do this in aligning itself with recent standard setting efforts in mathematics, for example.

Still, it will take a long time for the nation to deliver the educational performance envisioned by the standards. Some members of our subcommittee argued that NAEP should immediately reflect the content standards on the grounds that it makes no sense to gradually calibrate the thermometer once we know that 98.6 degrees is what is expected. But extreme

---

* Personal communication with Robert Glaser, Chair of the National Academy panel commissioned by NCES on the evaluation of the trial state NAEP.

differences between current practice and the content standards, and between practice and what is measured by NAEP, would strain credibility. After much discussion, we concluded that NAEP should both report results against the agreed upon content standards and report results of what is actually going on in the nation's classrooms. Presenting both kinds of information would recognize that a major departure from current practice will take time, but would also keep visible the gap between practice and expectation while tracking progress toward the goals.

Standard setting efforts, and the communication of those efforts, should move very quickly. The more people involved in these standards setting exercises the better. For NAEP, that means expanding the consensus building process to include thousands rather than hundreds. For other standard setting efforts such as NCTM, the message is that more people in schools and out need to be informed of the solid work that has been done.

2. Reporting NAEP results at state and national levels

There were only 7500 copies of the national goals report in the initial printing. A recent poll from the National PTA found that few parents knew about the goals or could name even one of them. NAEP suffers from the same problem. A national assessment expert urged one of our committee members to pick ten first rate teachers and ask them what NAEP stands for. None of them will know, he said. And sadly, he was right.

NAEP is not intended to drive classroom practice directly. That is a task better performed by local and state level assessments. But is it important that teachers and other citizens become at least as aware of NAEP reports as they are of the Dow Jones Index or the current unemployment figures. It is not sufficient to point out that this assessment is intended primarily to support the policy process. Without much wider familiarity with NAEP results, national and state policy leaders simply will not have the public or professional support they need to act on the results.

There are other reasons for expanding knowledge of NAEP. Teachers and students provide the data for NAEP. Individual schools receive no results that are of immediate use to them. This fact has often been cited to justify the view that the assessment understates the nation's educational performance because there is no incentive to perform on the test. Wider knowledge of NAEP would not create personal incentives but could in the context of heightened involvement in reform efforts in general give teachers and students some sense that they were participating in significant activities. As more teachers take leading roles in the design of state and local performance assessments, they will expect much greater information about NAEP.

Timing the release of NAEP results each year is important. We want maximum visibility for the NAEP results to focus attention on the goals. Release during the winter would be in the midst of budget negotiations at state and local levels. Test data appearing

then will enliven one side of the debate or the other but wont drive planning. Release in early summer will find most of the people who can do something with the results distracted by vacations, recovering from the year just completed, or engrossed in professional development programs planned much earlier. Release during the early fall will be lost in the activity that is familiar to anyone who has been in school.

The best time for release is in late August. This is when the nation is gearing up for another school year. During the months prior to release, educators, other partners in educational policy, and the media can be informed about what is coming so that people have time to coordinate their efforts with the anticipated reports. For example, states and schools can gain more visibility for their own performance reports in the fall if they are ready to comment on the most recent NAEP data.

And since NAEP results are and will remain such a critical part of the system to track the national education goals, it is particularly important that the National Governing Board and the National Goals Panel coordinate the release of their reports. Such coordination can only intensify the public attention given to both the NAEP results and the annual report of the Goals Panel.

When NAEP results appear, there should be a really massive communication strategy so that every teacher, every parent has the chance to read at least the basic findings. NAGB can contribute to a greater awareness by coordinating with other national assessment groups such as SAT and ACT so that a connected series of messages hits at almost the same time. The objective should be a powerful national message, which is then underscored by searching, thoughtful discussion in a hundred thousand arenas close to home. We would like to see every state, every local school board, every community, every dinner table discussion turn at some point to performance results.

3. Quality and breadth of items and tasks in each subject matter area

The pool of items for NAEP should be greatly expanded. The pool is much smaller than is commonly recognized. A wider pool of questions will give the NAEP tests more stability over the long term.

Expansion of the pool would also permit the annual release of a much larger number of items for public discussion without compromising the security of the tests. This will permit state assessments to include NAEP items and thus strengthen the connections among state assessments and national standards. Public discussion of released test items is one important way to focus attention on what performance is expected. Such discussion is commonplace after national assessments in some European nations where the press recounts public reaction to interesting test questions.

Performance based tasks are emerging now in great variety from teachers who are connected to the experiments with alternative forms of assessment. NAGB's attention to this work could contribute significantly to the quality of these tasks.

4. Subject matter and grade levels

NAEP should not be limited to the core subjects mentioned in the National Education Goals. Citizenship, the arts and foreign languages, for example, are also vital areas where the public needs reliable measures of performance. All of these parts of the curriculum are important, so how should the Goals Panel and NAGB decide which items to add to the assessment next? One reasonable approach would be to turn to those subject areas that are farthest along in laying the groundwork for assessment. The arts education community, for example, seems to have developed a solid basis for portfolio and performance assessment and should be considered an early candidate for NAEP assessment.

In general, the subcommittee is content with the subject matter and grade levels of the current NAEP approach. Some members note that while the subject area disciplines predominate in the National Goals and in most current assessment, there is growing interest in interdisciplinary approaches to curriculum and instructional practices. NAGB would do well to continue to track this interest and support research into assessment along these lines. We would also like to see research on the usefulness of assessments at the 12th grade because we are concerned about the low motivation for performance that appears characteristic of the end of high school today.

5. Achievement levels

The setting of achievement levels was the most controversial aspect of the 1990 state-by-state NAEP. But beyond the controversy is a fundamental point. The nation needs unambiguous, carefully defined achievement levels in all the core subject areas.

When results appear, the public needs to know what the achievement levels mean. The best definitions go beyond simple explanations to include examples and illustrations.

Setting the achievement levels cannot be hurried. It will take years to get it right, and probably the efforts of a very great many people. Technical adequacy will take relatively few individuals, but we are also aiming for widespread support, and that demands a very open process.

6. Performance tasks

Performance assessment has attracted so much attention that many people seem to embrace it without being troubled by the lack of a shared definition of what it means.

Performance assessments are those in which the "answer" is the behavior itself, rather than a description of it. For example, rather than identifying correct usage from among a number of alternatives, the student must actually write and demonstrate correct usage. Or instead of identifying correct statements about a particular art form, the student choreographs and performs the dance.

There are many forms of performance assessment, but all of them involve judgements about student work in relation to an objective set of criteria. Portfolios are collections of student work, often including rough drafts as well as the finished product. Portfolios represent work collected over an extended period, sometimes over more than one year. Performances or exhibitions refer to a demonstration of competence, either before a panel or a larger audience. Projects are a third form that involve the preparation and presentation of many kinds of information as in a research report or some other medium.

Performance assessment lacks the nine decades of experience and psychometrics that undergird more traditional forms of assessment, but one who has encountered the best of what is available in performance assessment would find it hard to doubt its power to motivate teaching and learning.

It is a strongly held belief among many educators, not yet supported by abundant research, although that is coming, that performance assessment gives a better picture of what students know and can do than does multiple choice testing. NAEP has for much of its history been the source of innovation in assessment. On the matter of performance assessment, NAEP is following now rather than leading. However, NAEP only has to turn to its own history for examples of leadership in performance assessment, such as its assessments in music and the visual arts in 1978–79.

We believe that NAEP should do as much as it can to join and encourage current experiments and development in performance assessment. For reasons of cost, it may appear that the only opportunities available to NAEP right now are open ended questions, writing samples, and opportunities for students to show their work on test booklets. But with many states now committed to full scale development of portfolio assessments among other measures, there is reason to consider the use of performance assessment as part of NAEP on a sample basis.

And the nation again needs NAEP leadership as it expands what is possible in performance assessment. NAEP's influence is such that if it turns in this direction, it will empower inventors elsewhere. But the influence can work in both directions. If performance assessment gains in practice, NAEP will lose credibility if it does not reflect practices that many teachers find sensible.

7. How often to assess?

There should be enough time between reports for schools and policy leaders to do something with the results and be able to conclude that the next test will reflect their efforts. This argues for a relatively long cycle between repetitions of say, a science assessment. Assessment on a national scale is complex, and it won't be done well if the assessing organization must continually gear up, then slow down. This argues for a steady state in which the annual assessment cycle is the same. The subjects tested can change each year, but there should be something being tested in much the same way each year.

The national goals tied to the year 2000 are the immediate driving force behind all of this, and it seems reasonable to have at least two data points on each of the core subject areas between now and the end of the decade. This argues for a relatively short cycle.

Put all of this together, and we conclude that a three-year cycle is the best option. Each of the core subjects would be tested every three years, with some part of the core tested and reported each year. Other subjects not part of the core would be assessed less frequently. The results of this effort would be included in the Goals Panel Report in the year following data collection.

The following table illustrates this proposed data collection schedule. The entries for the core subjects are: Rdng= Reading, Math= Mathematics, Wrtng= Writing, Scie=Science, Hstry= History, and Geog= Geography. Sub A through D represent possible non-core subjects. As reported in the table, 1992 NAEP data are currently being collected in Reading, Writing, and Mathematics.

PROPOSED NAEP SCHEDULE

1992	1993	1994	1995	1996	1997	1998	1999	2000
Rdng	Scie	Geog	Rdng	Scie	Geog	Rdng	Scie	Geog
Math	Hstry	Wrtng	Math	Hstry	Wrtng	Math	Hstry	Wrtng
Wrtng	Sub A	Sub B	Sub C	Sub D	Sub A	Sub B	Sub C	Sub D

8. Other statistics

There are two opposite views on how much data to collect concerning instructional practices and student characteristics. One position is that these data are powerful motivators for change because they allow state policy leaders to focus local attention on practices associated with poor performance. One example is the link between high TV viewing time and low academic performance. From this perspective, it is important to place the academic

performance data in context. Results are what count, but so is having an equitable chance to reach those high levels of performance.

The other position is that data about instructional practices should be largely removed from NAEP because the relationships between practice and results are very complex and reliable reporting would require an impossibly large array of data. Those who take this position point out that the burden on teachers and schools is already too high, and that they learned nothing new from this part of the assessment. It follows from this view that it would be better to drop most of the instructional practices data from NAEP and develop a separate research study on a sample basis.

There is obvious room for middle ground here, since no one disputes the need for valid results, for more information, for collecting only what can be reported, and for amplifying NAEP reports with additional scholarly studies on a sample basis. We would stress solid research to follow up themes that appear interesting, and would reduce the emphasis on simple cross tabulations in the reporting of NAEP results because they lead to inappropriate causal inferences. We recognize that this recommendation will require a vigorous research program and a budget to support it.

A final word

For most of the decade, NAEP will be the primary source of information on the nation's progress toward national education goals three and four. NAEP will not be the only source – many other elements of a national examination system are appearing and must be supported – and NAEP will not be a fixed system. NAEP is changing to incorporate national content standards, faster reporting, more performance based questions, and state by state reporting. The direction of these changes, when coupled with NAEP's two decades of solid reporting on educational progress, is encouraging.

But NAEP cannot be taken for granted. It must be reauthorized periodically, and this year is one such occasion. The state by state NAEP in particular is a temporary enhancement to the system, and while we think it should become a permanent feature of the national report card, it will disappear without Congressional action. The National Goals Panel can be the champion for this action.

REFERENCES

Ad Hoc Committee on the Relationship of NAEP to a National Examination System (Sept. 1991). *Summary of committee position statements* (mimeo). Washington, DC: National Assessment Governing Board.

Ad Hoc Committee on the Relationship of NAEP to a National Examination System (Oct. 1991). *Committee position statements* (mimeo). Washington, DC: National Assessment Governing Board.

Bock, R.D. & McCabe, P. (1990). *Toward a timely state NAEP* (report to the National Assessment Governing Board). Washington, DC: National Assessment Governing Board.

Elliott, E. (1991). *National testing: The gap between technicians and policy makers* (memo). Washington, DC: National Center for Education Statistics.

Elliott, E. (1991). *An individual student version of NAEP* (mimeo). Washington, DC: National Center for Education Statistics.

Ferrara, S.F. & Thornton, S.J. (1988). Using NAEP for interstate comparisons: The beginnings of a "national achievement test" and "national curriculum." *Educational Evaluation and Policy Analysis, 10(3),* 200–210.

Forsyth, R. A. (1991). Effects of item order and context on estimation of NAEP reading proficiency. *Educational Measurement: Issues and Practice, 10(3),* 10–16.

Haertel, E.H., Phillips, G.W., & Koretz, D.M. (April 1991). Should the National Assessment of Educational Progress be used to compare states? *Educational Researcher,* 17–21.

Harris, K.H. & Longstreet, W.S. (July/Aug. 1990). Alternative testing and the national agenda for control. *The Social Studies,* 90–93.

Irwin, P.M., Stedman, J.B. & Riddle, W.C. (1991). *National Assessment of Educational Progress: Background and status.* Washington, DC: Congressional Resource Service, the Library of Congress.

Madaus, G. (April 1991). *Testimony before the Select Education Subcommittee, Committee on Education and Labor.* Washington, DC: U.S. House of Representatives, Office of the Clerk, Office of Official Reports.

National Academy of Education (1991). *April 1991 interim report on the evaluation of the 1990 NAEP trial state assessment.* Washington, DC: Author.

National Assessment Governing Board (1991). *Response to the draft summative evaluation report on the National Assessment Governing Board's inaugural effort to set achievement levels on the National Assessment of Educational Progress.* Washington, DC: Author.

National Council on Education Standards and Testing. *Toward high standards: A report to Congress, the Secretary of Education, and the National Education Goals Panel (December 1991, draft).* Washington, DC: Author.

National Education Goals Panel (March 1991). *Measuring progress toward the national education goals: Potential indicators and measurement strategies.* Washington, DC: Author.

National Education Goals Panel (Sept. 1991). *Potential strategies for long-term indicator development.* Washington, DC: Author.

Phillips, G.W. & Finn, C.E., Jr. (April 1990). State-by-state comparisons can benefit education. *Educational Leadership,* 43-55.

Resnick, D.P. & Resnick, L.B. (April 1988). Understanding achievement and acting to produce it: Some recommendations for the NAEP. *Phi Delta Kappan,* 576-579.

Stufflebeam, D.L., R. Jaeger, and M. Scriven (1991). *Summative Evaluation of the National Assessment Governing Board's inaugural effort to set achievement levels on the National Assessment of Educational Progress* (draft report to the National Assessment Governing Board). Kalamazoo, MI: The Evaluation Center, Western Michigan University.

Truby, R.E. & Philips, G.W. (Feb. 1990). State-by-state assessments will yield new insights. *The School Administrator,* 63-66.

NATIONAL EDUCATION GOALS PANEL

INTERIM EXECUTIVE DIRECTOR

Martin Orland

PROFESSIONAL STAFF

Nancy Delasos
Edward Fuentes
Laura Lancaster
Leslie Lawrence
Cynthia Prince
Charles Walter
Emily Wurtz

SUPPORT STAFF

Tia Cosey
Edna Wilson

BEST COPY AVAILABLE

1850 M Street, NW Suite 270 Washington, DC 20036
(202) 632-0952 FAX (202) 632-0957